

WHAT IS CLAIMED IS:

1 1. A multi-computer system, comprising a plurality of local nodes
2 interconnected by a shared memory, each local node comprising
3 a local processor,
4 a local memory,
5 a local communications protocol stack, and
6 a shared memory interface system operable to provide a local shared memory
7 network between the local nodes, and a global shared memory network between the
8 local nodes and one or more remote nodes by capturing packets from the local
9 communications protocol stacks and routing the captured packets over the shared
10 memory.

1 2. The multi-computer system of claim 1, wherein the shared memory
2 interface system on each local node comprises a local shared memory virtual adapter
3 and a global shared memory virtual adapter;

4 the local shared memory virtual adapters being operable to capture locally
5 addressed packets from the local communications protocol stacks and to route the
6 captured packets for physical transport over the shared memory; and

1 the global shared memory virtual adapters being operable to capture globally
2 addressed packets from the local communications protocol stacks and to route the
3 captured packets for physical transport over the shared memory.

1 3. The multi-computer system of claim 2, wherein the local shared
2 memory virtual adapters appear to the local communications protocol stacks as
3 device drivers for physical network adapters connected to the local shared memory
4 network, and the global shared memory virtual adapters appear to the local
5 communications protocol stacks as device drivers for physical network adapters
6 connected to the global shared memory network.

1 4. The multi-computer system of claim 2, wherein the global shared
2 memory virtual adapters are responsive to a common global address for the global
3 shared memory network.

1 5. The multi-computer system of claim 4, wherein the global shared
2 memory virtual adapters are operable to capture from the local communications
3 protocol stacks packets destined to a global network address and transmitted from a
4 local network address.

1 6. The multi-computer system of claim 5, wherein the global shared
2 memory virtual adapters are operable to route in-bound packets to other local nodes
3 over the global shared memory network.

1 7. The multi-computer system of claim 6, wherein the shared memory
2 interface system on each local node is operable to maintain in local memory a data
3 structure identifying active local nodes connected to the global shared memory
4 network.

1 8. The multi-computer system of claim 1, wherein one or more local
2 nodes comprise one or more physical network adapters for connection to one or
3 more remote nodes.

1 9. The multi-computer system of claim 8, wherein the shared memory
2 interface system is operable to route packets to local nodes over the global shared
3 memory network in accordance with an open shortest path first (OSPF) routing
4 protocol.

1 10. The multi-computer system of claim 9, wherein local nodes comprising
2 physical network adapters are configured as OSPF area border routers.

1 11. The multi-computer system of claim 10, wherein packets are routed
2 over the global shared memory network preferentially to local nodes configured as
3 OSPF area border routers.

1 12. The multi-computer system of claim 10, wherein OSPF cost metrics are
2 set so that routes to OSPF area border routers are preferentially over the local shared
3 memory network.

1 13. The multi-computer system of claim 1, wherein the shared memory
2 interface system on each local node supports multicast and broadcast transmissions
3 over the shared memory for the local shared memory network and the global shared
4 memory network.

1 14. The multi-computer system of claim 13, wherein a broadcast ring
2 structure and a multicast ring structure are allocated in shared memory for each of
3 the local and global shared memory networks.

1 15. The multi-computer system of claim 14, wherein the broadcast ring
2 structure and the multicast ring structure are reallocated to an active node in
3 response to a failure of a local node originally allocating the broadcast ring structure
4 or the multicast ring structure.

1 16. The multi-computer system of claim 1, wherein for each of the local
2 and global shared memory networks a pair of transmit/receive ring structures are
3 allocated in shared memory for each pair of local nodes.

1 17. The multi-computer system of claim 16, wherein each transmit/receive
2 ring structure corresponds to a pre-allocated number of fixed size scribble buffers in
3 shared memory.

1 18. The multi-computer system of claim 16, wherein the shared memory
2 interface system on each local node is operable to allocate a transmit/receive ring
3 structure in shared memory for each of the other local nodes.

1 19. The multi-computer system of claim 18, wherein the shared memory
2 interface system on each local node is operable to connect to a transmit/receive ring
3 structure allocated by a given node in response to receipt of a broadcast packet from
4 the given node.

1 20. The multi-computer system of claim 19, wherein the shared memory
2 interface system on each local node is operable to allocate a transmit/receive ring

3 structure for the given node in response to receipt of the broadcast packet from the
4 given node.

1 21. The multi-computer system of claim 16, wherein a read pointer and a
2 write pointer are associated with each transmit/receive ring structure.

1 22. The multi-computer system of claim 21, wherein a write pointer is
2 modifiable only by a transmitting node and a read pointer is modifiable only by a
3 receiving node.

1 23. The multi-computer system of claim 1, wherein the shared memory is
2 implemented by a global shared memory facility, a distributed shared memory
3 facility, or a logically shared memory facility.

1 24. A computer program residing on a computer-readable medium in a
2 multi-computer system comprising a plurality of local nodes interconnected by a
3 shared memory, each local node comprising a local processor, a local memory, and a
4 local communications protocol stack, the computer program comprising computer-
5 readable instructions for causing a computer to:

6 provide a local shared memory network between the local nodes, and a global
7 shared memory network between the local nodes and one or more remote nodes by
8 capturing packets from the local communications protocol stacks and routing the
9 captured packets over the shared memory.

1 25. The computer program of claim 24, wherein the computer program
2 comprises computer-readable instructions for causing a computer to capture from the
3 local communications protocol stacks packets destined to a global network address
4 and transmitted from a local network address.

1 26. The computer program of claim 25, wherein the computer program
2 comprises computer-readable instructions for causing a computer to route in-bound
3 packets to other local nodes over the global shared memory network.

1 27. The computer program of claim 26, wherein the computer program
2 comprises computer-readable instructions for causing a computer to maintain in local
3 memory a data structure identifying active local nodes connected to the global shared
4 memory network.

1 28. The computer program of claim 27, wherein the computer program
2 comprises computer-readable instructions for causing a computer to route packets to
3 local nodes over the global shared memory network in accordance with an open
4 shortest path first (OSPF) routing protocol.

1 29. The computer program of claim 28, wherein the computer program
2 comprises computer-readable instructions for causing a computer to configure local
3 nodes comprising physical network adapters as OSPF area border routers.

1 30. The computer program of claim 29, wherein the computer program
2 comprises computer-readable instructions for causing a computer to route packets
3 over the global shared memory network preferentially to local nodes configured as
4 OSPF area border routers.

1 31. The computer program of claim 24, wherein the computer program
2 comprises computer-readable instructions for causing a computer to allocate in
3 shared memory a broadcast ring structure and a multicast ring structure for each of
4 the local and global shared memory networks.

1 32. The computer program of claim 31, wherein the computer program
2 comprises computer-readable instructions for causing a computer to reallocate the
3 broadcast ring structure and the multicast ring structure to an active node in
4 response to a failure of a local node originally allocating the broadcast ring structure
5 or the multicast ring structure.

1 33. The computer program of claim 24, wherein, for each of the local and
2 global shared memory networks, the computer program comprises computer-
3 readable instructions for causing a computer to allocate in shared memory a pair of
4 transmit/receive ring structures for each pair of local nodes.

1 34. The computer program of claim 33, wherein each transmit/receive ring
2 structure corresponds to a pre-allocated number of fixed size scribble buffers in
3 shared memory.

1 35. The computer program of claim 33, wherein the computer program
2 comprises computer-readable instructions for causing a computer to allocate a
3 transmit/receive ring structure in shared memory for each of the other local nodes.

1 36. The computer program of claim 35, wherein the computer program
2 comprises computer-readable instructions for causing a computer to connect to a
3 transmit/receive ring structure allocated by a given node in response to receipt of a
4 broadcast packet from the given node.

1 37. The computer program of claim 36, wherein the computer program
2 comprises computer-readable instructions for causing a computer to allocate a
3 transmit/receive ring structure for the given node in response to receipt of the
4 broadcast packet from the given node.

1 38. A method of processing packets in a multi-computer system comprising
2 a plurality of local nodes interconnected by a shared memory, each local node
3 comprising a local processor, a local memory, and a local communications protocol
4 stack, the method comprising:

5 providing a local shared memory network between the local nodes, and a
6 global shared memory network between the local nodes and one or more remote
7 nodes by capturing packets from the local communications protocol stacks and
8 routing the captured packets over the shared memory.